Why aren't Users Using Protection? Investigating the Usability of Smartphone Locking

Why aren't Users Using Protection? Investigating the Usability of Smartphone Locking

Nicholas Micallef*, Mike Just^, Lynne Baillie^, Martin Halvey[+], Hilmi Güneş Kayacik[‡]
* Glasgow Caledonian University, Glasgow, UK, nicholasmicallef@gmail.com
^ Heriot-Watt University, Edinburgh, UK, {m.just, l.baillie}@hw.ac.uk
[+] University of Strathclyde, Glasgow, UK, martin.halvey@strath.ac.uk
[‡] FICO, 181 Metro Dr. San Jose CA, guneskayacik@fico.com

ABSTRACT
One of the main reasons why smartphone users do not adopt screen locking mechanisms is due to the inefficiency of entering a PIN/pattern each time they use their phone. To address this problem we designed a context-sensitive screen locking application which asked participants to enter a PIN/pattern only when necessary, and evaluated its impact on efficiency and satisfaction. Both groups of participants, who prior to the study either locked or did not lock their phone, adopted our application and felt that unlocking their phone only when necessary was more efficient, did not annoy them and offered a reasonable level of security. Participants responded positively to the option of choosing when a PIN/pattern is required in different contexts. Therefore, we recommend that designers of smartphone locking mechanisms should consider ceding a reasonable level of control over security settings to users to increase adoption and convenience, while keeping smartphones reasonably secure.

Author Keywords
Usability; Mobile HCI; User Experience; Adoption; Authentication

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User Interfaces – Input devices and strategies, evaluation.

INTRODUCTION
Current smartphone screen locking mechanisms are not as efficient as they could be because every time a user wishes to use their phone (e.g. make a call, open an app etc.) they need to enter a PIN, pattern or password. This additional step may be the reason why 64% of smartphone users do not lock their phones [4]. Of the users that do use a locking mechanism, 40-47% find it "annoying" [5,8,10]. This suggests that users who regard convenience to be important do not adopt locking mechanisms and those users who give high importance to security adopt locking mechanisms but still feel annoyed when using them. This leads to two usability problems: efficiency and satisfaction.

Attempts to improve smartphone authentication have typically considered alternative mechanisms, such as graphics [3], touch biometrics [12,20] and gestures [13,18], though such solutions have not adequately addressed efficiency and satisfaction [19]. To date context-dependent solutions [6,9] have relied upon location sensors alone for determining when to require an explicit unlock with a locking mechanism. Using location alone, however, introduces insider attacks (attacks carried out by people who have unrestricted access to a victim's space), which are increasingly becoming a major concern for users [5,16]. In our approach we augment location sensors with environment-related sensors (i.e., noise, light, magnetic field and accelerometer) to characterize the environmental surroundings of the smartphone. For example, only when the phone detects a change in environmental surroundings (e.g., different noise levels due to having friends over) would the phone ask for explicit unlocking. This adds an extra layer of security against insider attacks over location sensors only solutions. Thus, a context-dependent solution that uses location with environment sensors has two aims (1) to increase efficiency by reducing the instances in which users need to enter an explicit unlocking mechanism to only those in which the changes in the usage pattern and environment make it necessary to input a PIN/pattern, (2) to improve security for those users who currently do not lock their phone while maintaining a reasonable level of security for those users who currently lock their phone. Therefore, we hypothesize that this type of solution could increase use among traditional non-adopters (users who do not lock their phone), and reduce the level of annoyance to current adopters, whilst still providing an acceptable level of convenience and security.

Thus, we designed and implemented a context-sensitive screen locking application and then evaluated its adoption using a 3-phase user study. We began by constructing an environment profile (Phase I) using 1 week of sensor data [11]. In the second week (Phase II) the application changes from passive data collection and analysis to an active screen locking mechanism which determines when to prompt the participants for a PIN/pattern depending upon the match of

the current sensor readings to the environment profile. In the third week of the study (Phase III) we evaluate the level of adoption of our application by giving the participants the choice of using our screen locking application or not using it at all and reverting to their previous screen locking mechanism (PIN/Pattern/No Lock) in five different contexts ("Home", "Work", "Other Places", "On the move" and "New Places"). We are not aware of any other research which has tried to evaluate a screen locking mechanism by involving users in such an active and direct manner. Our objectives were:

O1: To understand how users perceive our screen locking application in terms of annoyance, satisfaction and security.

O2: To understand if prompting users with a PIN/pattern only when necessary will increase the adoption rate of our screen locking application.

O3: To identify which kind of users (from those who currently do not lock their phone, and those who lock it) would adopt our screen locking application.

O4: To understand the contexts in which users would adopt our screen locking application.

In the following sections we describe the security threat model of this mechanism, provide an overview of the algorithm used in our screen locking application and explain the methodology used to fulfill our objectives. We then present our empirical, perception and adoption results and discuss how these results address our objectives.

RELATED WORK

There are two main areas of related work relevant to this research that we present in this section: context-sensitive screen locking solutions, and studies of user perception of smartphone locking mechanisms.

Context-sensitive screen locking solutions

Attempts to improve smartphone screen locking have typically considered alternative mechanisms, such as graphics [3], touch biometrics [12,20] and gestures[13,18], though such solutions have not adequately addressed the issues of efficiency and satisfaction [19]. Several recent solutions have focused on context-sensitive screen locking mechanisms. For instance, Gupta et al. [6] describe a context profiler which uses location traces to detect places of interest for a user, and they use the Bluetooth and Wi-Fi sensors to estimate a place's familiarity. The calculated level of safety for a location is used to determine which kind of unlock mechanism is shown to the user, e.g. to request a PIN or not. Unfortunately, they do not involve users to confirm whether the familiarity of the place is correct. In our study, we confirm that our algorithms are correctly identifying locations by asking users to confirm the home and work locations. We use algorithms which aggregate related locations [14] to minimize user involvement and ask users to confirm home and work locations, only rather than confirming all five locations (home, work, other places, on the move and new places) to not overly annoy users.

Hayashi et al. [9] use location data (e.g. Wi-Fi MAC address, IP address etc.) to modulate the level of authentication required. They provide the user with either a weak (none or PIN) or strong (PIN or password) unlocking mechanism based on location. With their framework they reduced explicit unlocks by 68%. They conducted two user studies. In the first they claim that they did not obtain positive usability results as they used a combination of PIN or password as the strong locking mechanism. In the second they used only the PIN as the strong locking mechanism, resulting in better usability results since users found the use of a password in the first study to be "too much of a burden" [9]. Riva et al. [17] used a variety of contextual factors to estimate the probability that the legitimate user is in close proximity to their device, from which they determined the mechanism to present in the unlock screen. Their study reduced the number of explicit unlocks by 42%. The main limitation of Riva et al.'s research is that it was not evaluated in the field.

Our research builds on previous work in this area [6,9,17] in two ways: (1) we use environment-related sensors in conjunction with location-based sensors in order to add an extra layer of security over a location-only-sensor solution, (2) the user involvement in our field study, and particularly the third week of our study that evaluates the level of adoption of our mechanism (which no other research has evaluated in detail). Since one aim of our research is to increase the adoption of screen locking mechanisms, in the next section we present empirical research that investigated users' perceptions of locking mechanisms.

Locking perceptions

Von Zezschwitz et al. compared the performance of Android-like patterns to PINs [21]. They found that for input speed and error rate, PINs performed better than patterns. But for ease-of-use, feedback and likeability, users preferred pattern locks. Despite these results, such explicit locking methods are still not considered efficient [10,22]. Van Bruggen et al. studied the use of "interventions" to improve the adoption of smartphone locking methods among Android users [2]. They tried to tackle the "bring your own device" to work problem by evaluating several types of intervention messages (based on incentives, morality and deterrence) with the aim of convincing users to adopt a locking mechanism on their phone. However, they did not obtain a significant increase in the adoption of locking mechanisms, and they concluded that such interventions are not worth the cost.

Harbach et al. [8] carried out an online survey of locking behaviors and risk perceptions, followed by a month-long experience sampling experiment in which participants were constantly asked to report on shoulder surfing while unlocking their phones. They found that participants who

do not lock their phone were very satisfied with their choice and indicated very few situations where they would rather have locked their screen. For participants that did lock their devices, dissatisfaction was quite moderate in public situations, as participants valued protection slightly more in that context. They recommended that researchers should focus on decreasing the number of unlocks by deploying context dependent locking mechanisms. We took this recommendation into consideration for our study.

Egelman et al. [5] study several threat models by comparing participants' perceptions of the sensitivity level of the data stored on their smartphones. They found that 1/4 of participants locked their devices due to the possibility of insider attacks from friends and family (also highlighted by [16]). For this reason they recommend that the decision on whether the screen locking mechanism is shown or not should go beyond context and usage of the phone. This concern is addressed in two ways in our study, firstly our context-sensitive screen locking application determines when to show the PIN/pattern based on the characteristics of the environmental surroundings of the smartphone and secondly the users themselves select which contexts they want to be prompted for a PIN/pattern. The empirical research discussed in [2,5,8] provided the motivation required to evaluate the usability of our screen locking application. We describe the security threat model of our application in the next section.

SECURITY THREAT MODEL

Context-sensitive solutions which only use location sensors improve usability (over the use of an explicit locking mechanism) by reducing the number of explicit unlocks required, but this makes them vulnerable to insider attacks [5]. The aim of this work is to build a mechanism which increases the adoption of screen locking mechanisms, hence we had to cater for both users who lock and do not lock their phone. Thus, we could not only focus on improving usability but we also had to obtain a reasonable level of security. As reported upon in previous work [11,15] sensor-driven authentication mechanisms are capable of detecting uninformed and informed security attacks in a reasonable amount of time and at the same time keep a low percentage of false positives. Hence, our main focus is on evaluating the usability of a mechanism that augments location sensors with environment sensor readings. Such a mechanism improves the usability for those users who lock their phone with an explicit mechanism by reducing the number of times that they have to enter a PIN/pattern, and maintains a reasonable level of security by asking for an explicit unlock only when the sensor readings do not match the environmental profile. In the case of those users who do not lock their phone this mechanism improves security by using a locking mechanism, which asks for an explicit unlock when the environmental sensor readings do not match the environmental profile. At the same time it keeps a reasonable level of usability since they are not asked to explicitly unlock their phone each time they need to use it. Although we are trying to handle insider attacks by augmenting location with environmental sensors we still consider these types of attacks to be challenging and we do not claim that our mechanism can catch all occurrences of these attacks. For this reason when there is a concern for such attacks, say at work, we gave the user the option (in Phase III) to adopt full PIN/pattern protection in those chosen contexts. Figure 1 illustrates how easy it is to select the locking mechanism for each context.

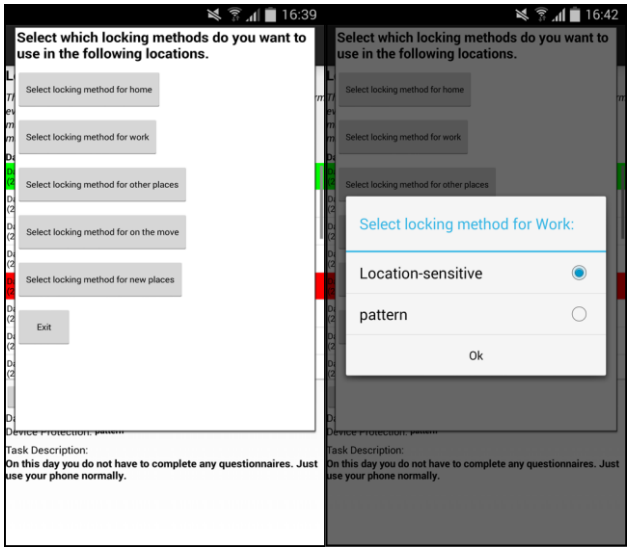

Figure 1. Screen shots of application.

ALGORITHM OVERVIEW

In this section we outline the algorithm used to build our application. The algorithm combines five different sensors with the aim of characterizing many of the features that make up the surrounding environment of the smartphone. Every 1 min the application samples the following sensors: the accelerometer (x, y & z coordinates for 10 sec), Wi-Fi (access point names and IDs), light (light level for 5 sec), microphone (ambient noise levels for 10 sec) & magnetometer (x, y & z values for magnetic field for 10 sec). These configurations are based on the individual sensor requirements tested in previously conducted pilot studies. These features are aggregated into one vector which consists of the following attributes for each feature (with the exception of Wi-Fi which is not numeric, and therefore we use the names and IDs): mean, mode, median, stdev, min, max and range.

The decision to use these five sensors and a 1 min sampling rate is based on a trade-off between the level of security and battery consumption. To have an optimal level of security to identify short transitioning events (when users move from one room to another) we required a sampling rate of 1 min or less (see Figure 2): in pilot tests that we carried out before this study we found that as we decreased the sampling frequency the attack detection time increased considerably. We considered a 30 sec sampling rate for a

higher level of security, but due to battery consumption issues we would have had to use fewer sensors. In our pilot studies, we found that the battery of a Samsung Galaxy S4 using a 1 min sampling rate and these five sensors (with the settings listed above) took 26 hrs to drain for normal phone usage (instead of 29 hrs without our application) and 10 hrs for high phone usage (instead of 10.4 hrs without our app) – see Figure 2. To avoid increasing the study participants' burden we opted not to increase the sampling rate further. On the other hand, a less frequent sampling rate (e.g. 5 mins) could have included more sensors (e.g. GPS, proximity) but this would have increased the attack opportunity.

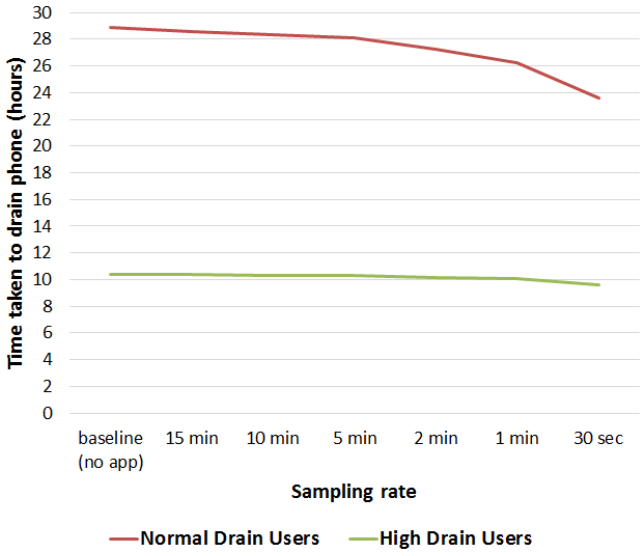

Figure 2. Time taken to drain phone per sampling rate for normal and high drain users.

When building the profile (in Phase I), environment sensor data was sampled and J48 decision trees were established for each location. Before building the ambient profiles using J48 machine learning models the sensor data had to undergo two processing stages. In the first processing stage the raw sensor data was converted into sensor attributes (mean, mode, median, etc) as defined above. The second processing stage calculated a location anchor to attach to the sensor attributes created in the first stage. The location anchor was required because to define J48 machine learning models (ambient profiles) which learned the location from the provided sensor attributes, the sensor attributes had to be used as the input and the location anchors had to be used as the outputs of the machine learning algorithms. A specific module was implemented using Weka's Java component to create these ambient profiles locally (on the phone) on the early morning of the 8th day (the first day of Phase II).

When later validating against the profile (in Phases II & III) at a particular location, current sensor readings were compared to the decision tree conditions in the profile for that location. For example, for location X, the profile expected a noise level less than 40dB and a light level greater than 100lux. If sensor readings were not within the set of conditions for that particular location or the current location was not recognized by the profile ("new places"), then PIN/pattern entry was requested. There can be multiple sensor configurations for each location (e.g. at home the decision tree contained a set of conditions for the morning and a completely different set for the night).

We use the smartphone location to define different contexts for our study from which we confirmed actual location (ground truth) information from our participants. Thus, in Phase I of our study (see next section: Overall Approach) participants were asked to identify the Wi-Fi access point names that related to their home and work contexts. From these two contexts, we further defined the "on the move" contexts (continuously changing Wi-Fi access points) and the "other" contexts (a "catch all" for all other common locations, e.g. coffee shop or pub, which are not continuously changing and are not home or work). In Phases II and III we also use the "New places" context to define all those locations which were not visited when creating the profile during Phase I.

USER STUDY

In our user study we categorize two user groups; those who currently do not lock their phone, and those who lock it. In this way we can understand how our application would affect these two groups of users who have different security requirements. Hence, we asked participants to complete perception questionnaires with statements about annoyance, convenience and security to determine how they perceive our application in terms of these three properties throughout the three phases of the study (addressing O1). The user study included a third phase in which participants decide whether to adopt our screen locking mechanism or rollback to their previous mechanism based on context, to try to understand if prompting users with a lock screen only when necessary will increase the adoption rate (addressing O2). We also wanted to understand whether either of the two groups of users is more likely to adopt our application (addressing O3). This extra phase is also used to understand the contexts in which participants are more likely to adopt our application (addressing O4). Prior to our study we obtained ethical approval from our University's Ethics Committee.

Overall Approach

To address our research objectives we conducted an exploratory user study consisting of three phases, and each phase took one week to complete (see Table 1). We constructed the profile with one week of data based on research conducted by Kayacik et al. [11], which found that in most cases a sensor-driven profile stabilizes after a week. During the installation participants chose their current smartphone screen locking mechanism, which they used during Phase I. We offered three options (no screen lock,

PIN, and pattern) as these represent the most used screen locking mechanisms [4] and we recruited participants that used each of these mechanisms. For participants who chose either PIN or pattern, their "original" mechanism was also used for explicit screen locking in Phases II and III. For participants who indicated that they do not lock their phone, their "original" mechanism was "no lock", and they were asked to select either a PIN or a pattern screen locking method to be used for explicit screen locking during Phases II and III. In Table 1 we provide an outline of all phases.

Phase	Steps involved	Duration
Setup	• Study explained to participants. • Demographic questionnaire completed. • Installation of application.	30 min
Phase I: Data collection	• Collected sensor data to create context-sensitive profile. • Participants used their original screen locking mechanism and evaluated its usability.	1 week (20 min for usability feedback)
Phase II: Evaluated our app	• Participants evaluated the usability of our application. • Participants completed usability questionnaires at the end of each day. • Participants completed an end of phase questionnaire.	1 week (20 min for usability feedback - end of each day +20 min for end of phase feedback)
Phase III: Evaluated adoption of our app	• Participants decided whether and in what contexts they wanted to adopt our app: home, work, other places, on the move, new places. • Participants evaluated the usability of our application in the adopted contexts. • On the 4th day participants were given the option to change setup and start/stop using our application in different contexts. • Participants completed an end of study questionnaire.	1 week (20 min for usability feedback - end of each day + 20 min at the end of study feedback)

Table 1. Main procedure, steps and duration of study.

We evaluated the usability of the original mechanism (Phase I), our screen locking application (Phase II) and the adoption of our application (Phase III) by using the original System Usability Scale (SUS) [1] and a perception questionnaire adapted from [8]. In the questionnaire we asked for responses to the following four statements (at the end of Phase I and everyday during Phases II & III), with answers on a 5-point Likert scale (1-strongly disagree to 5-strongly agree).

S1: The number of times in which I had to unlock my phone today was annoying.

S2: I felt secure with today's phone protection mechanism.

S3: Overall, the number of times in which I unlocked the phone today was convenient.

S4: I wish there was a more convenient way of unlocking my phone.

These questionnaires allowed us to evaluate the impact on usability of being prompted to unlock the phone with our screen locking application (Phase II) more/less than their original mechanism (Phase I). We used the end-of-study questionnaire to understand whether participants would use our screen locking application and how it ranked in the spectrum of current screen locking mechanisms in terms of annoyance, convenience and security. We also collected logging data to compute the number of times in which participants unlocked their phone using the PIN/pattern and how long the participants took to unlock their phone.

Participants

Participants were recruited through social media and word of mouth. Participants that completed all three phases of the study were compensated with a £30 Amazon voucher. The compensation was directly related to the completion of the study and not whether participants adopted our application in Phase III. So as not to influence participants' choice of locking mechanism, we only met with participants at the beginning and end of the study and we did not have any contact with them when they were transitioning from Phase II to Phase III, or when the application asked them whether they want to change their setup in the middle of Phase III. The pre-requisites to participate in this study were to own a fairly recent android phone (with hardware capabilities similar to a Samsung Galaxy S3), to have a regular home/work routine for the three weeks of the study and to have regularly used one of the following three smartphone screen locking mechanisms: no screen lock, PIN or pattern.

Participants were divided into two groups by their use of an original screen locking mechanism: Group 1: "No Lock" - participants who use a swipe (default Android setting); Group 2: "Lock" - participants who use either a PIN or pattern. We grouped the PIN and pattern together as it was already challenging to recruit such participants since the vast majority of the population does not lock their phone (64% according to [4]). For this reason we also did not include password participants. Consumer reports [4] confirm that 23% of smartphone users use a 4-digit PIN and only 13% use a longer PIN, password or a pattern.

Overall, we recruited 25 participants over 2.5 months, with 20 completing the study. All participants completed the three phases of the study in the same order (see Table 1). Using the logging data we confirm that none of the 20 participants uninstalled the application before the end of the study. Due to constraints placed by our university ethics committee we were required to uninstall the application from the participants' phones after these three weeks. Of the five that did not complete, for two participants the data was not complete (i.e. phone was not collecting sensor data when in sleep mode), while the other three participants dropped out due to unexpected work/personal

commitments. The participants of the study used a variety of phones: Samsung Galaxy S3 (7), Samsung Galaxy S4 (7), Sony Xperia S1 compact (3), Samsung Galaxy Note 3 (1) & LG Nexus 4 (2). The average participant age was 32 (22-61), med=30. We opted to use the participants' personal phones, rather than giving them a smartphone, as we wanted to encourage them to continue their normal behaviour throughout the study.

RESULTS

In this section we present the empirical, perception and adoption results collected across the three phases of the study in order to evaluate the usability and adoption of a context-sensitive screen locking application which augments location with environment-related sensors. Since the perception results were not normally distributed we analyzed them using Friedman's test to understand the effect of using locking mechanisms across the study's three phases. We used this test to also understand the difference in the perception of annoyance, convenience and security when compared to different locking methods. Wilcoxon pair-wise comparisons were used to determine which combinations were statistically significant. Note that in Phase III, two participants returned to their original mechanism and did not adopt our application (see Adoption Results for further details). Hence, in all Phase III results reported in the empirical results (Tables 2 & 3) and ratings of statements (Table 4) we exclude any ratings from these 2 participants (N=18). When ranking locking mechanisms according to perceived annoyance, convenience and security we use the rankings of all 20 participants (N=20).

Empirical Results

Table 2 shows the average number of times (per day and according to context) in which participants entered a PIN/pattern to unlock their phone during the three phases of the study together with the average number of times that the participants activated their phone during a particular phase. Participants were always asked to unlock their phone with a PIN/pattern when they were in new places because if a place was not visited when the profile was created (Phase I) then it was not trusted. Our counts excluded those instances where the screen was "on" but locked such as when checking the time or checking for new notifications. In Phase II "Lock" participants unlocked their phone using a PIN/pattern an average of 16 out of 56 activations per day. Phase III results in Table 2 do not include any unlocks from the two participants that did not adopt our application, but it includes unlocks of those participants who adopted our application in at least 1 context. The "No Lock" group experienced a drop in unlocks from 34% in Phase II to 24% in Phase III due to most participants in this group reverting back to using no lock when at home and at work (see Adoption Results for further details).

Table 3 shows the average aggregate time taken (in sec) to enter PIN/pattern (per day) during the three phases of the study. Friedman statistical analysis of these results across the three phases did not highlight any significant differences. In the discussion section we compare these empirical results to the perception and adoption results to understand how they impacted the adoption results and perception of annoyance, convenience and security across the three phases.

Group	Context	Phase I (N=20)	Phase II (N=20)	Phase III (N=18)
"No Lock"	Home	0 of 25	7 of 26	3 of 23
	Work	0 of 15	6 of 21	3 of 18
	Other Places	0 of 13	2 of 6	2 of 8
	On the move	0 of 9	1 of 8	1 of 5
	New Places	0 of 0	7 of 7	5 of 5
	Overall	**0 of 62 (0%)**	**23 of 68 (34%)**	**14 of 59 (24%)**
"Lock"	Home	19 of 19	6 of 23	4 of 19
	Work	13 of 13	3 of 16	3 of 15
	Other Places	9 of 9	2 of 8	2 of 6
	On the move	4 of 4	1 of 5	1 of 4
	New Places	0 of 0	4 of 4	2 of 2
	Overall	**45 of 45 (100%)**	**16 of 56 (29%)**	**12 of 46 (26%)**

Table 2. Average number of times that the participants entered a PIN/pattern per day to unlock their phone.

Group	Phase I (N=20)	Phase II (N=20)	Phase III (N=18)
"No Lock"	0 (0) σ =0	131 (103) σ =94	86 (55) σ =74
"Lock"	240 (202) σ =260	105 (103) σ =99	90 (67) σ =81

Table 3. Average time taken per day (sec) to enter PIN/pattern- mean, (median) & standard deviation (σ).

User perception of the application

We evaluated perception by analyzing the average of the daily results obtained in Phases II and III, as well as the results collected at the end of Phase I. The target was to understand whether participants changed their opinion regarding the different mechanisms used across these three phases. In Table 4 we list the mean, median (in parentheses) and standard deviation (σ) results.

Friedman statistical analysis of the results of statement S1 (see Overall Approach Section) across the 3 phases did not highlight any significant differences. This might be related to the fact that the perception of annoyance of the "No Lock" group was relatively consistent across the three phases. This means that across the entire study the "No Lock" participants disagreed that the number of unlocks was "annoying". On the other hand, "Lock" participants had a decline (not significant) in annoyance and standard deviation (see Table 4) when comparing the use of our application (in Phases II & III) to their original mechanism. Hence both groups disagreed that the number of times in which they had to unlock their phone was annoying.

The results for S2 (see Table 4) show that the "No Lock" participants neither agree nor disagree about feeling secure with their original screen locking mechanism, while "Lock" participants seem to agree that they feel secure with their original locking mechanism. Statistical analysis of the results of statement S2 across the three phases did not highlight any significant differences. This is related to the fact that the "Lock" group had a consistent feeling of security across all three phases of the study. On the other hand the perception of feeling secure for the "No Lock" group improved (not significant) and even had a reduction in standard deviation after using our application in Phases II & III. Thus there was a trend (albeit not significant) which showed that when using our screen locking application the "No Lock" group felt more secure.

	Group	Phase I (N=20)	Phase II (N=20)	Phase III (N=18)
S1	"No Lock"	2 (2) **σ** =0.87	2.62 (2.08) **σ** =0.90	2.02 (2) **σ** =0.41
	"Lock"	3.13 (3.5) **σ** =1.3	1.89 (2) **σ** =0.60	1.79 (2) **σ** =0.39
S2	"No Lock"	3.22 (3) **σ** =0.83	3.6 (4) **σ** =0.66	4 (4) **σ** =0.5
	"Lock"	3.63 (4) **σ** =0.71	3.68 (4) **σ** =0.58	3.86 (4) **σ** =0.12
S3	"No Lock"	3.77 (4) **σ** =0.83	3.26 (3.56) **σ** =0.81	3.81 (4) **σ** =0.53
	"Lock"	3 (3) **σ** =1.3	4.02 (4) **σ** =0.26	3.82 (4) **σ** =0.57
S4	"No Lock"	3.44 (4) **σ** =1.01	3.03 (3) **σ** =0.72	2.83 (3) **σ** =0.83
	"Lock"	3.50 (4) **σ** =1.02	2.57 (2.21) **σ** =0.99	2.68(2.64) **σ** =0.77

Table 4. Average ratings for statements S1-S4 for each of the three phases. (1. Strongly disagree, 5. Strongly agree).

The standard deviation results for S3 (see Table 4) show that there was quite a large spread (**σ**=1.3) in the convenience of the "Lock" participants that had to unlock their phone using their original mechanism (Phase I). Statistical analysis of the results of statement S3 across the three phases did not highlight any significant differences. Despite finding no significance, the results in Table 4 still show a trend (albeit not significant) that "Lock" participants found the number of times in which they had to unlock their phone using our application (in Phases II & III) to be more convenient than their original mechanism.

Regarding S4 the results show that when they were using their original locking mechanism (Phase I) participants from both groups seemed to be either neutral or agree about whether there should be a more convenient way of unlocking their phone. Friedman statistical analysis of the results across the three phases found significant differences (X^2 (2) =6.377, p = 0.041). Wilcoxon pair-wise comparisons across the results of the different phases (adjusted Bonferroni p value =0.017) returned significant differences when comparing the results from the original mechanism to the use of our application in Phase II (z=-2,437, p=0.015) and in Phase III (z=-1.159, p=0.008). These results suggest that as participants from both groups used our application, their opinion shifted significantly from agreeing that they wished for a more convenient way of unlocking their phone to a more neutral/disagree stance. When we analyzed the "Lock" group results (for S1-S4) we did not find any significant differences between PIN and Pattern users.

	Annoyance	Convenience	Security
Password	1.84 (1) **σ** =1.21	4.47 (5) **σ** =0.90	2 (2) **σ** =0.94
PIN	2.47 (2) **σ** =1.22	3.84 (4) **σ** =1.01	2.17 (2) **σ** =0.90
Pattern	2.31 (2) **σ** =0.82	3.58 (4) **σ** =1.07	2.9 (3) **σ** =1.10
No Lock	4.42 (5) **σ** =1.30	1.55 (1) **σ** =1.26	5 (5) **σ** =0
Our App	3.95 (4) **σ** =0.91	2.42 (2) **σ** =1.07	2.79 (3) **σ** =1.40

Table 5. Perception of annoyance (1=most annoying, 5=least annoying), convenience (1=most convenient, 5=least convenient) & security (1=most secure, 5=least secure) - mean, (median) & standard deviation (σ).

At the end of Phase III we asked all 20 participants to rank the screen locking mechanisms used in this study (and Passwords) according to perceived annoyance, convenience and security. We asked participants to rank passwords as well because we believed that they had the appropriate experience of using passwords. Our motivation was to understand where participants would place our application in the spectrum of current screen locking mechanisms with respect to these three properties. Table 5 shows that with respect to annoyance, participants from both groups ranked our application to be the second least annoying out of the listed screen locking mechanisms, with only "no lock" being less annoying. Friedman statistical analysis across annoyance rankings of these five locking mechanisms found significant differences (X^2 (4) =38.331, p < 0.001). Wilcoxon pair-wise comparisons across the results of the different locking mechanisms (adjusted Bonferroni p value = 0.005) showed significant differences when comparing the results from our application with "PIN" (z=-3.022, p=0.003), "Password" (z=-3.440, p= 0.001) and "Pattern" (z=-3.675, p<0.001). For "no lock" the difference with our application was not significant.

Table 5 shows that with respect to convenience, participants from both groups ranked our application to be the second most convenient mechanism out of the listed current locking mechanisms, with only "no lock" being more convenient. Statistical analysis across convenience rankings of these five locking mechanisms found significant differences (X^2 (4) =40.080, p < 0.001). Pair-wise comparisons across the results of the different locking mechanisms showed significant differences when comparing the results of our application to the "PIN" (z=-2.917, p=0.004) and

"Password" (z=-3.348, p=0.001). While in the case of the "no lock" and "pattern" locking mechanisms the differences were not significant.

Table 5 shows that overall, with respect to security, participants ranked our application to be the third most secure mechanism when compared to the current locking mechanisms, with "no lock" and "pattern" being less secure. Statistical analysis across security rankings of these five locking mechanisms found significant differences (X^2 (4) = 48.012, $p < 0.001$). Pair-wise comparisons across the results of the different locking mechanisms showed significant differences when comparing the results of our application to the "no lock" (z=-3.769, $p < 0.001$). When compared to the "PIN", "Password" and "Pattern" our application did not have a significant difference for security perception.

We also asked participants to evaluate usability using the original SUS during all the three phases of the study (end of Phase I, daily for Phases II & III). Overall the participants' original mechanism (Phase I) and our application (Phase II) received the same SUS Score: 74% (SD=±12.2516) and (SD=±6.3508) respectively. When adopted in Phase III our application received a SUS Score of 80% (SD=±7.6887) with a 90% confidence level of (90-93.54). There is no statistically significant difference between the overall SUS scores obtained across the three phases of the study, but these results suggest a SUS score for our approach no worse than current screen locking choices.

Adoption results

When moving from Phase II to Phase III, 18 out of 20 participants chose to adopt our application in at least one of these five contexts: "home", "work", "other places", "on the move" and "new places". More than five "No Lock" participants adopted our application in the following three contexts: "other places", "on the move" and "new places". Also, all "Lock" participants adopted our application "at home". The two participants that did not adopt our application were both in the "No Lock" group. The first participant explained his choice by saying that he is not ready to trade convenience with any kind of increase in security. The other participant explained his choice by saying that when he is at home or at work he does not feel that the people around him are a threat and when he is on the move, in other places or in new places he always keeps his phone in his pocket, therefore he does not need to use any smartphone locking mechanisms. Only two of the remaining 18 participants (both "No Lock" participants) changed the configuration in the middle of Phase III. One participant decided to stop using our application at work because he felt that he no longer needed security in this context but continued using it in other places, on the move and in new places. The other participant initially selected to use our application in other places, on the move and in new places but did not select it for home & work. But, after three days of not having any lock at work he realized that he actually wanted more security, so he also started using it at work.

Group	Over all	Home	Work	Other Places	On the Move	New Places
"No Lock"	8	1	4	7	5	8
"Lock"	9	9	5	3	4	2

Table 6. Final adoption results distributed by context.

At the end of the study participants were asked whether they would use our application if it was commercially available and in which contexts they would adopt it. Overall 17 participants said that they would adopt it (see "Overall" column in Table 6). This means that only one participant who opted to use our application in Phase III decided that he would not subsequently adopt it. This participant was in the "Lock" group and he explained his choice by saying that he is not ready to trade a higher level of convenience with a decreased level of security, therefore he prefers to have full security all the time. The 17 other participants said that they would keep the same setup that they selected in Phase III.

DISCUSSION

In this section we reflect on our research objectives by discussing the results outlined in the previous section.

Is it secure, less annoying and convenient? *(O1)*

The participants that originally did not lock their phone ("No Lock") and those that locked their phone with either a PIN or pattern ("Lock") felt secure when using our screen locking application in Phases II & III of the study. Inputting a PIN/pattern only when necessary made the "No Lock" group improve their security perception with respect to when they were not inputting any kind of lock. In the case of the "Lock" group the considerable decrease in number of unlocks that they experienced while using our application did not translate into a significant decrease in security perception, meaning that our application still made this group feel reasonably secure. These results were confirmed when we asked the participants to rank locking mechanisms according to security perception. When comparing the security perception of our application to the current locking mechanisms, the participants ranked our application significantly better than the no lock but relatively similar to the PIN and pattern (see Table 5). This is quite an important finding since in terms of security perception these results place our application on a par with industry standards such as the PIN and the pattern. This is just a user perception result and we are not claiming that our mechanism is more secure than current screen locking mechanisms.

Despite introducing an average of 22 unlocks per day (see Table 2) and an average time spent entering a PIN/pattern of 131 seconds per day (see Table 3), the participants in the "No Lock" group still felt that our screen locking

application was not annoying and they rated our application slightly higher than when they were not using a lock. "Lock" participants experienced a considerable but not significant decrease in annoyance when comparing their original mechanism to our application. This suggests that the decrease in time spent to enter a PIN/pattern from Phase I to Phase II (see Table 3) by an average of 135 seconds per day and the decrease in the number of times in which they needed to enter the PIN/pattern from 100% (45 of 45) to an average of 29% (16 of 56) per day (see Table 2) seems to have contributed to this shift in perceived annoyance. These results were confirmed when we asked the participants to rank locking mechanisms according to annoyance. When comparing the annoyance perception of our application to the current screen locking mechanisms, the participants ranked it to be on the same level as the no lock mechanism and significantly better than the PIN and pattern locking methods (see Table 5). Thus, in terms of annoyance perception these results place our application on a better level than current locking mechanisms.

From a convenience perspective, although the "No Lock" participants experienced an increase in unlocks and time spent to enter a PIN/pattern they still felt that our screen locking application (in both Phases II & III) was convenient and rated the convenience level slightly less (not significant) than their original mechanism. Despite experiencing a sharp decrease in the number of times that they needed to unlock their phone and time spent to enter PIN/pattern (see Tables 2 and 3), participants in the "Lock" group did not experience a significant change in opinion when comparing the convenience of their original mechanism to our application (in both Phases II & III). Although, they still seemed to agree that the number of times that they unlocked the phone was convenient. This lack of a significant difference can be attributed to the fact that some of the participants still considered the PIN/pattern to be relatively convenient. However, when comparing the rankings of the perceived convenience of our application to the current screen locking mechanisms, participants ranked it similarly to the no lock and the pattern but significantly better than the PIN (see Table 5). Thus, in terms of convenience participants placed our application on a better level than the PIN and on the same level as the pattern.

Is it adoptable? And by whom? *(O2 & O3)*

18 out of 20 participants adopted our screen locking application in at least one context during Phase III and only one of these 18 participants said that he would not use our screen locking application if it was commercialized. This gives an overall adoption rate of 85% (17/20). This high adoption rate we believe can be explained by the results in Phase III (see Table 4) where both groups still ranked our application in a similar manner (sometimes even better) as they did in Phase II. This shows that participants of the study responded positively to the idea of choosing when a PIN/pattern is required in different contexts. The convenience results of S4 (see Table 4) also justify the high adoption rate because after only two weeks of using our application participants in both groups already exhibited a significant shift (from agree to quite close to disagree) in their opinion of whether they still required a more convenient way of unlocking their phone. The high adoption rate is also confirmed by the fact that the two groups of participants evaluated this mechanism similarly in all metrics (Table 4). There is no instance (for both Phases II & III) where one group evaluated our application significantly different from the other group. Similarly we found no instance where PIN and Pattern users evaluated our application significantly different from each other.

This means that participants who currently do not lock their phone adopted our application because they felt that entering the PIN/pattern only when necessary was convenient, did not annoy them and offered them a consistent level of security. On the other hand, participants who lock their phone, adopted our application because it was more efficient (since it reduced the number of times they had to enter their PIN/pattern), less annoying, more convenient and they still felt secure. Hence, one of the major findings of our study is that both participant groups seem to indicate that a screen locking application which augments location with environment-related sensors has all the necessary qualities for being adopted.

In which contexts would it be adopted? *(O4)*

Our study confirmed Harbach et al.'s finding [8] that smartphone users were mostly dissatisfied with locking mechanisms when they were in private spaces. In fact, all "Lock" participants adopted our screen locking application at "home", since they had to explicitly unlock their phone during Phase II an average 6 times per day out of 23 unlocks (see Table 2), thereby choosing the convenience of fewer PIN/pattern unlocks. In the questionnaires that we collected at the end of Phases II & III, participants stated that they did not consider "home" to be a major security threat. Despite "No Lock" participants having to explicitly unlock their phone during Phase II an average of 7 times per day out of 26 unlocks when at "home" and an average of 6 times per day out of 21 unlocks when at "work" (see Table 2), several "No Lock" participants preferred to use no lock at all in these contexts. In most cases the participants in this group (see Table 6) did not consider their home or work to be a threat therefore they did not want to enter any PIN/pattern when in these contexts.

We can also report that most "No Lock" participants adopted our application in "other places", "on the move" and "new places" because they felt that these are the contexts in which their phone is more likely to be exposed to a security threat. Most of the participants in the "Lock" group felt that the "other places" and "on the move" contexts were more likely to be exposed to a security threat. For this reason they selected the PIN/pattern in these contexts. This shows that in both groups there were

common trends in the selection of contexts in which the participants decided to adopt our screen locking application. When linking these results to the results obtained in Phase III (see Tables 2-4) we can conclude that the one-size-fits-all design currently being provided by manufacturers and recommended by security experts is one of the main reasons why smartphone users are not adopting current screen locking mechanisms.

LIMITATIONS AND FUTURE WORK

The results obtained by this study are quite promising and the setup used was sufficient to address our objectives. However, due to the limitations of having a limited amount of participants (N=20) using one smartphone platform, future work should extend this study to a larger user-base and implement the application on additional smartphone platforms (IOS and windows phone) to provide further confirmation of our findings. In the larger study we plan to include users who use biometric screen locking mechanisms such as fingerprint and face recognition. While these are not currently widely used they may be at some stage in the future. Also, participants of this study were users with a regular daily routine. An interesting improvement would be to include users who do not have a regular daily routine so that we would evaluate how these kinds of users would perceive the use of our application.

CONCLUSION

Most smartphone users are not adopting screen locking mechanisms [4], while those who do adopt them find them annoying [5,8,10]. Through a user study we tried to tackle this problem by evaluating a context-sensitive screen locking application which used location and environment-related sensors to ask participants to enter a PIN/pattern only when necessary, thus improving efficiency and hence usability.

We found that 85% of all participants reported that they would adopt our screen locking application if it was available on their phone. This shows that asking users to unlock their phone only when necessary can increase the adoption rate of screen locking mechanisms. In terms of security perception, participants placed our application on a par with industry standards such as the PIN and the pattern. As regards to annoyance perception, participants thought that our application was better than the current locking mechanisms. With respect to convenience perception, participants placed our application on a better level than the PIN and on the same level as the pattern. Hence, participants perceived our application to be reasonably secure and at the same time it did not have the inconveniences of the existing screen locking mechanisms.

We found that both groups of participants would adopt our application in a similar manner because there were no instances in which the results of the two groups were statistically different. Both groups of participants felt that unlocking the phone only when necessary was convenient, did not annoy them and made them feel secure. This means that despite having divergent security concerns both groups adopted our context-sensitive screen locking application. Hence demonstrating that the one-size-fits all design used at the moment does not meet the needs of smartphones users. Therefore, we recommend that designers of smartphone authentication methods should consider ceding a reasonable level of control over security settings to users to increase adoption and convenience of authentication. Ultimately this would make smartphones and the sensitive data stored on them more secure since it would increase the use of security mechanisms among traditional non-adopters.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no PIIF-GA-2011-301536.

REFERENCES

1. Brooke, J. SUS: A quick and dirty usability scale. In *Usability Evaluation in Industry*. 1996, 198–194.
2. Van Bruggen, D., Liu, S., Kajzer, M., Striegel, A., Crowell, C.R., and D'Arcy, J. Modifying smartphone user locking behavior. *Proceedings of SOUPS '13*, ACM Press.
3. Chiang, H.-Y. and Chiasson, S. Improving user authentication on mobile devices: a touchscreen graphical password. *Proceedings of MobileHCI '13*, ACM Press, 251–260.
4. Cell Phone Security. Wireless Threats - Consumer Reports. 2013. http://consumerreports.org/privacy0613.
5. Egelman, S., Jain, S., Portnoff, R.S., Liao, K., Consolvo, S., and Wagner, D. Are You Ready to Lock ? Understanding User Motivations for Smartphone Locking Behaviors. *Proceedings of CCS '14*, ACM Press, 750–761.
6. Gupta, A., Miettinen, M., Asokan, N., and Nagy, M. Intuitive Security Policy Configuration in Mobile Devices Using Context Profiling. *International Conference on Privacy, Security, Risk and Trust*, IEEE (2012), 471–480.
7. Hall, M., Frank, E., and Holmes, G. The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, (2009), 10–18.
8. Harbach, M., von Zezschwitz, Emanuel, Fichtner, A., De Luca, A., and Smith, M. It's a hard lock life: A field study of smartphone (un) locking behavior and risk perception. *Proceedings of SOUPS '14*, USENIX.
9. Hayashi, E., Das, S., Amini, S., Hong, J., and Oakley, I. CASA: context-aware scalable authentication. *Proceedings of SOUPS '13*, ACM Press.

10. Jakobsson, M., Shi, E., Golle, P., and Chow, R. Implicit authentication for mobile devices. *Proceedings of HotSec'09,* USENIX.

11. Kayacık, H.G., Just, M., Baillie, L., Aspinall, D., and Micallef, N. Data Driven Authentication: On the Effectiveness of User Behaviour Modelling with Mobile Device Sensors. *Proceedings of MoST 2014*.

12. De Luca, A., Hang, A., Brudy, F., Lindner, C., and Hussmann, H. Touch me once and i know it's you! Implicit Authentication based on Touch Screen Patterns. *Proceedings of CHI '12*, ACM Press, 987–996.

13. De Luca, A., von Zezschwitz, E., Nguyen, N.D.H., et al. Back-of-device authentication on smartphones. *Proceedings of CHI '13*, ACM Press, 2389–2398.

14. Micallef, N., Just, M., Baillie, L., and Kayacik, H.G. Stop questioning me! Towards optimizing user involvement during data collection on mobile devices. *Proceedings of MobileHCI '13*, ACM Press, 588–593.

15. Micallef, N., Kayacik, H.G., Just, M., Baillie, L., and Aspinall, D. Sensor Use and Usefulness: Trade-Offs for Data-Driven Authentication on Mobile Devices. *Proceedings of IEEE PerCom 2015*.

16. Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., and Beznosov, K. Know your enemy: the risk of unauthorized access in smartphones by insiders. *Proceedings of MobileHCI '13*, ACM Press, 271–280.

17. Riva, O., Qin, C., Strauss, K., and Lymberopoulos, D. Progressive Authentication: Deciding When to Authenticate on Mobile Phones. *Proceedings of 21st USENIX Security Symposium*, (2012).

18. Sae-Bae, N., Ahmed, K., Isbister, K., and Memon, N. Biometric-rich gestures: A novel approach to authentication on multi-touch devices. *Proceedings of CHI '12*, ACM Press, 977–986.

19. Schlöglhofer, R. and Sametinger, J. Secure and usable authentication on mobile devices. *Proceedings of MoMM '12*, ACM Press, 257–262.

20. Xu, H., Zhou, Y., and Lyu, M. Towards Continuous and Passive Authentication via Touch Biometrics: An Experimental Study on Smartphones. *Proceedings of SOUPS '14*, ACM Press 187–198.

21. Von Zezschwitz, E., Dunphy, P., and De Luca, A. Patterns in the wild: a field study of the usability of pattern and pin-based authentication on mobile devices. *Proceedings of MobileHCI '13*, ACM Press, 261–270.

22. Von Zezschwitz, E., De Luca, A., and Hussmann, H. Honey, I shrunk the keys: influences of mobile devices on password composition and authentication performance. *Proceedings of NordiCHI '14*, ACM Press, 461–470.